


```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5936362"
/clone_1ib="NIH MGC 106"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: blood; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC Library."
BASE COUNT      221 a      296 c      288 g      197 t
ORIGIN
Query Match      93.8%; Score 734.6; DB 14; Length 1002;
Best Local Similarity 98.3%; Pred. No. 1.5e-177;
Matches 774; Conservative 0; Mismatches 9; Indels 4; Gaps 3;

QY      1  ATGGGAAGTCGCGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTCTCTGTC 60
DB      96  ATGGGAAGTCGCGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTCTCTGTC 155
QY      61  CAAGGCCAGGAGCTGTGACCATGAAAGCAGAGAAAGCCAAAGCCGAGCCCTG 120
DB      156  CAAGGCCAGGAGCTGTGACCATGAAAGCAGAGAAAGCCAAAGCCGAGCCCTG 215
QY      121  GGCAAGTTCCCGGAGTGGCCGCGGAGCAGTGTGAGAGCTCGGGAGGCAATTGACC 180
DB      216  GGCAAGTTCCCGGAGTGGCCGCGGAGCAGTGTGAGAGCTCGGGAGGCAATTGACC 275
QY      181  ATGCTCTCTGAGAGATGAGATGCTGTGACGCTGTCTGAAGTCTCAGGCAAGAGTAT 240
DB      276  ATGCTCTCTGAGAGATGAGATGCTGTGACGCTGTCTGAAGTCTCAGGCAAGAGTAT 335
QY      241  AACATCCCGACGCTCCAGCGTGGCCAAAGTCTCCCAATGGGTGCTGATAGAGGCGTGGC 300
DB      336  AACATCCCGACGCTCCAGCGTGGCCAAAGTCTCCCAATGGGTGCTGATAGAGGCGTGGC 395
QY      301  AGGAGAAAGCAGAGAACTGCTGTGTAACCTGTGAAACCTGTGAGGGGCTTCTCTATC 360
DB      396  AGGAGAAAGCAGAGAACTGCTGTGTAACCTGTGAAACCTGTGAGGGGCTTCTCTATC 455
QY      361  CGGAGAGACCAAGCAGAGAGAGCTTTACTCTCTGTGAGTCCGCTCAGCCGCGCTTGA 420
DB      456  CGGAGAGACCAAGCAGAGAGAGCTTTACTCTCTGTGAGTCCGCTCAGCCGCGCTTGA 515
QY      421  TCTGGGAGCGGATCAGACACTACAGATCCATGCTTGTGCAATGCTGAGCTGTATATC 480
DB      516  TCTGGGAGCGGATCAGACACTACAGATCCATGCTTGTGCAATGCTGAGCTGTATATC 575
QY      481  TCACCGGCGCTCACTTCCCTCACTCAGAGCCCTGTGTGAGCACTTACTCTGAGTGTGCG 540
DB      576  TCACCGGCGCTCACTTCCCTCACTCAGAGCCCTGTGTGAGCACTTACTCTGAGTGTGCG 635
QY      541  GATGACATCTGCTGCTTACTCAAGAGGCTGTGTCTGTGAGAGGCTGTGCGCTCTCT 600
DB      636  GATGACATCTGCTGCTTACTCAAGAGGCTGTGTCTGTGAGAGGCTGTGCGCTCTCT 695
QY      601  GGCAAGAGATATACCCCTACTGTGACTGTGAGAGAGAACCACTCAACTGAGAAAGAGTG 660
DB      696  GGCAAGAGATATACCCCTACTGTGACTGTGAGAGAGAACCACTCAACTGAGAAAGAGTG 755
QY      661  GACAGCTCCCTCTCTGTTTCTGAAGCTGCAACA-GGGAGAGATCTTCTTCTAGTGAAGG 719
DB      756  GACAGCTCCCTCTCTGTTTCTGAAGCTGCAACA-GGGAGAGATCTTCTTCTAGTGAAGG 815
QY      720  --TCTCCGGAAGTCCCTCAGCTTCAATCAGCTGAGATGACGA-GGCTGTCTCTTTGGA 776
DB      816  GTCTCCGGGAGATCCCTCAGCTTCAATCAGCTGAGATGACGA-GGCTGTCTCTTTGGA 875

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5936362"
/clone_1ib="NIH MGC 106"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: blood; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC Library."
BASE COUNT      219 a      311 c      283 g      204 t      3 others
ORIGIN
Query Match      83.5%; Score 654; DB 14; Length 1020;
Best Local Similarity 95.4%; Pred. No. 6.6e-157;
Matches 748; Conservative 0; Mismatches 27; Indels 9; Gaps 7;

QY      1  ATGGGAAGTCGCGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTCTCTGTC 60
DB      212  ATGGGAAGTCGCGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTCTCTGTC 271
QY      61  CAAGGCCAGGAGCTGTGACCATGAAAGCAGAGAAAGCCAAAGCCGAGCCCTG 120
DB      272  CAAGGCCAGGAGCTGTGACCATGAAAGCAGAGAAAGCCAAAGCCGAGCCCTG 331
QY      121  GGCAAGTTCCCGGAGTGGCCGCGGAGCAGTGTGAGAGCTCGGGAGGCAATTGACC 180
DB      332  GGCAAGTTCCCGGAGTGGCCGCGGAGCAGTGTGAGAGCTCGGGAGGCAATTGACC 391
QY      181  ATGCTCTCTGAGAGATGAGATGCTGTGACGCTGTCTGAAGTCTCAGGCAAGAGTAT 240
DB      392  ATGCTCTCTGAGAGATGAGATGCTGTGACGCTGTCTGAAGTCTCAGGCAAGAGTAT 451
QY      241  AACATCCCGACGCTCCAGCGTGGCCAAAGTCTCCCAATGGGTGCTGATAGAGGCGTGGC 300

```

Db 452 AACATCCCGACGCTCCACGTCGCAAACTCTCCATGGGTGCTATAGAGGCTGAC 511
 Qy 301 AGGGAAGAAAGCAAGAACTGCTGTTTACCTGGAAACCTTGGAGGCTTCTCATC 360
 Db 512 AGGGAAGAAAGCAAGAACTGCTGTTTACCTGGAAACCTTGGAGGCTTCTCATC 571
 Qy 361 CGGAGAGCCAGACAGAGAGGCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCA 420
 Db 572 CGGAGAGCCAGACAGAGAGGCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCA 631
 Qy 421 TCTTGGAGCCGGATCAGACATCAAGATCCATGCTTGGCAATGCTGGCTGATCATC 480
 Db 632 TCTTGGAGCCGGATCAGACATCAAGATCCATGCTTGGCAATGCTGGCTGATCATC 691
 Qy 481 TCACCGGCTCAGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 539
 Db 692 TCACCGGCTCAGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 751
 Qy 540 GGATGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 598
 Db 752 GGATGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 811
 Qy 599 CTGGCAAGATATAT-CCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 657
 Db 812 CTGGCAAGATATATAT-CCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 871
 Qy 658 CTGGCAAGATATATAT-CCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 714
 Db 872 CTGGCAAGATATATAT-CCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 931
 Qy 715 GAGGG--TCTCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 771
 Db 932 GAGGGGCTTCTCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 991
 Qy 772 TTGG 775
 Db 992 TTGG 995

RESULT 3

B0052468

LOCUS

DEFINITION

AGENCOURT_6868422 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:593772

5' mRNA sequence.

ACCESSION

B0052468

VERSION

B0052468.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE

1 (bases 1 to 1069)

AUTHORS

NIH-MGC http://mgi.nci.nih.gov/

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cga@bbs.fda.gov

Tissue Procurement: Dr. Daniel McVicar, DBS/NCI

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC Clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L16M2118 row: n column: 13

High quality sequence stop: 681.

Location/Qualifiers

1. 1069

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:593772"

/clone_lib="NIH_MGC_106"

/issue_type="natural killer cells, cell line"

/lab host="DH10B (phage-resistant)"
 /note="Organ: Blood; Vector: pOTB7, Site 1: XhoI, Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCAAGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH MGC Library."

BASE COUNT

230 a 328 c 300 g 205 t 6 others

Query Match

Best Local Similarity 93.4%; Pred. No. 8.1e-154;

Matches 734; Conservative 0; Mismatches 0; Indels 52; Gaps 4;

Qy 1 ATGGGAGTCTGCGCCAGAGAAATCTGCGCAAGCCCAAGCTTGATCTCTGTC 60
 Db 89 ATGGGAGTCTGCGCCAGAGAAATCTGCGCAAGCCCAAGCTTGATCTCTGTC 148
 Qy 61 CAAGCCAGAGGACTGTGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
 Db 149 CAAGCCAGAGGACTGTGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 208
 Qy 121 GCGAGTTTCCGCGAGAGTGGCCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
 Db 209 GCGAGTTTCCGCGAGAGTGGCCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 268
 Qy 181 ATGCTCTGAGAGATGAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 240
 Db 269 ATGCTCTGAGAGATGAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 328
 Qy 241 AACATCCCAAGGCTCCAG 300
 Db 329 AACATCCCAAGGCTCCAG 388
 Qy 301 AGGGAAGAAAGCAAGAACTGCTGTTTACCTGGAAACCTTGGAGGAGCTTCTCATC 360
 Db 389 AGGGAAGAAAGCAAGAACTGCTGTTTACCTGGAAACCTTGGAGGAGCTTCTCATC 448
 Qy 361 CGGAGAGCCAGACAG 420
 Db 449 CGGAGAGCCAGACAG 508
 Qy 421 TCTTGGAGCCGGATCAGACATCAAGATCCATGCTTGGCAATGCTGGCTGATCATC 480
 Db 509 TCTTGGAGCCGGATCAGACATCAAGATCCATGCTTGGCAATGCTGGCTGATCATC 568
 Qy 481 TCACCGGCTCAGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 539
 Db 569 TCACCGGCTCAGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 623
 Qy 540 GGATGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 599
 Db 624 -----GGGCTGGCCGCTGCC 639
 Qy 600 TGGCAAGATATATAT-CCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 659
 Db 640 TGGCAAGATATATAT-CCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 699
 Qy 660 GGAAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 718
 Db 700 GGAAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 759
 Qy 719 GTCTCCGAGG-AGTCCCTGAGCTTGTATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 777
 Db 760 GTCTCCGAGGAGCTTGTATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 819
 Qy 778 GATGCC 783
 Db 820 GATGCC 825

RESULT 4
AL541041 877 bp mRNA linear EST 16-FEB-2001
LOCUS AL541041.LTI.FL002_PL1 Homo sapiens cDNA clone CS0DE05YK23 5 prime
DEFINITION mRNA sequence.
ACCESSION AL541041
VERSION AL541041.1 GI:12871733
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 877)
AUTHORS Li W.B., Gruber C., Jesse J., and Polayes D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr.
FEATURES
source
1. 877
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DE05YK23"
/clone_1b="LTI.FL002_PL1"
/lab_host="DH10B"
/note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand
cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-stranded cDNA was digested with Not I
and cloned into the Not I and Eco RV sites of the
pCMVSPORT 6 vector. Library was constructed by life
technologies. Contact: Feng liang, life technologies, a
division of invitrogen 9800 Medical Center Drive Rockville
, Maryland 20850, USA Fax: (1) 301 610 8371 Email:
fliang@lifetech.com URL:
http://fulllength.invitrogen.com"

BASE COUNT 192 a 262 c 251 g 170 t 2 others
ORIGIN

Query Match 62.4%; Score 488.6; DB 9; Length 877;
Best Local Similarity 99.4%; Pred. No. 1.5e-114;
Matches 488; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

1 ATGGGAAGTCTCCGAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 60
2 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
3 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
4 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
5 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
6 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
7 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
8 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
9 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
10 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
11 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
12 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
13 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
14 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
15 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
16 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
17 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
18 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
19 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
20 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
21 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
22 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
23 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
24 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
25 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
26 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
27 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
28 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
29 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
30 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
31 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
32 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
33 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
34 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
35 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
36 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
37 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
38 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
39 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
40 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
41 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
42 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
43 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
44 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
45 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
46 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
47 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
48 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
49 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
50 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
51 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
52 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
53 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
54 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
55 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
56 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
57 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
58 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
59 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
60 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
61 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
62 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
63 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
64 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
65 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
66 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
67 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
68 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
69 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
70 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
71 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
72 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
73 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
74 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
75 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
76 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
77 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
78 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
79 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
80 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
81 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
82 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
83 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
84 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
85 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
86 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
87 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
88 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
89 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
90 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
91 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
92 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
93 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
94 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
95 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
96 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
97 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
98 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
99 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446
100 ATGGGAAGTCTGCGCAGAGAAATATCTGCAAGCCCAAGCTTGAGTCTCTGTC 446

Db 807 TCCTGGACCGGATCAGACATCAGAGATCCAGCTTGAGTCTGCTGATC 866
Gy 481 TCACGGCGCT 491
Db 867 TCACGGCGCT 877

RESULT 5
AK020837 926 bp mRNA linear HTC 19-JAN-2002
LOCUS AK020837
DEFINITION Mus musculus adult retina cDNA, RIKEN full-length enriched library,
clone:A930009E21:similar to SRC-LIKE ADAPTER PROTEIN, full insert
sequence.
ACCESSION AK020837
VERSION AK020837.1 GI:12861542
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult retina cDNA to mRNA,
clone_1b:RIKEN full-length enriched mouse cDNA library
clone:A930009E21.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K.,
Itoh M., Kono H., Okazaki Y., Muramatsu M., and Hayashizaki Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 2
AUTHORS Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K.,
Itoh M., Kono H., Okazaki Y., Muramatsu M., and Hayashizaki Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 3
AUTHORS Kawai J., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
Kono H., Akiyama J., Nishi K., Katsunaka T., Taehiro H., Itoh M.,
Sun N., Ishii Y., Nakamura S., Hazama M., Nishie T., Harada A.,
Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
Yoneda Y., Inoue K., Togawa Y., Izawa M., Ohtsuka E., Watabiki M.,
Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsunaga S., Kawai J.,
Okazaki Y., Muramatsu M., Inoue Y., Kira A., and Hayashizaki Y.
TITLE RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multichannel sequencer
JOURNAL Nature Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
Aizawa K., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamashita I.,
Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
Kadota K., Matsuda H., Ashburner M., Baralov S., Casavant T.,
Pletschmann W., Gasterland T., Gissi C., King B., Kochina H.,
Kuenli P., Lewis S., Matsuo Y., Nishida T., Pesele G.,
Quackenbush J., Schmitt L.M., Stabili F., Suzuki R., Tomita M.,
Wagner J., Washio T., Sakai K., Okido T., Furuno M., Aono H.,
Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N.,
Carninci P., de Bonaldo M.F., Brownstein M.J., Bull C.,
Fletcher C., Fujita M., Gariboldi M., Gietlinch S., Hill D.,
Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P.,
Marchionni L., Mashima J., Mazzarelli J., Mombaeate P., Nordone P.,
Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H.,
Sato K., Schonbach C., Seta T., Shibata Y., Storch K.F., Suzuki H.,
Toyo-Oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohatsu S.,
and Hayashizaki Y.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409 (6821), 685-690 (2001)
MEDLINE 21085660

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: L1AM10418 row: C column: 07
 High quality sequence start: 2
 High quality sequence stop: 566.
 Location/Qualifiers

FEATURES

source

1..566

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4520382"
 /clone_lib="NIH_MGC_91"
 /tissue_type="adenoecarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: Prostate; Vector: PCWV-SPORT6; Site_1: NCI;
 Site_2: Sall; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.4 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH MGC library."

BASE COUNT

116 a 187 c 152 g 110 t 1 others

ORIGIN

Query Match 51.3%; Score 402; DB 12; Length 566;
 Best Local Similarity 100.0%; Pred. No. 1.9e-92;
 Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 382 GGCTTACTCTCTGTGAGTCCGCTCAGCCGCTGATCTCGGAGCCGATCAGAC 441
 Db 13 GGCTTACTCTCTGTGAGTCCGCTCAGCCGCTGATCTCGGAGCCGATCAGAC 72
 Qy 442 TACAGATCACTGCTTGAATGCTGCTGATCATCTGACCGGCTCACTTCCC 501
 Db 73 TACAGATCACTGCTTGAATGCTGCTGATCATCTGACCGGCTCACTTCCC 132
 Qy 502 TCACTCCAGGCTGCTGAGCACTTACTGAGCTGGCGGATGACATCTGCTACTC 561
 Db 133 TCACTCCAGGCTGCTGAGCACTTACTGAGCTGGCGGATGACATCTGCTACTC 192
 Qy 562 AAGAGCCCTGCTCTCTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 621
 Db 193 AAGAGCCCTGCTCTCTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 252
 Qy 622 GTGACTGTGAGAGCACTCACTCACTGAGAGGCTGAGAGGCTGAGAGGCTGAG 681
 Db 253 GTGACTGTGAGAGCACTCACTCACTGAGAGGCTGAGAGGCTGAGAGGCTGAG 312
 Qy 682 GAAGCTCCAGAGGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 741
 Db 313 GAAGCTCCAGAGGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 372
 Qy 742 TACATGAGCTGATGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 783
 Db 373 TACATGAGCTGATGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 414

RESULT 7

LOCUS B0054265 986 bp mRNA linear EST 29-MAR-2002
 DEFINITION AGENCOURT_6830248 NIH_MGC_106 Homo sapiens CDNA clone IMAGE:5936339
 5', mRNA sequence.
 ACCESSION B0054265
 VERSION B0054265.1 GI:19813605
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 986)
 NIH-MGC <http://imgc.ncbi.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Dr. Daniel McVicar, DMS/NCI
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: L1AM10418 row: C column: 12
 High quality sequence stop: 515.
 Location/Qualifiers

FEATURES

source

1..986

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5936339"
 /clone_lib="NIH_MGC_106"
 /tissue_type="natural killer cells, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: blood; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; CDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCAGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH MGC library."

BASE COUNT

211 a 291 c 276 g 207 t 1 others

ORIGIN

Query Match 49.3%; Score 386.2; DB 14; Length 986;
 Best Local Similarity 98.0%; Pred. No. 2.7e-88;
 Matches 444; Conservative 0; Mismatches 3; Indels 6; Gaps 5;

Qy 1 ATGGAGATCTGCGCAGAGAAATCTCTGCAAGCCCAAGTTAGTCTCTCTC 60
 Db 279 ATGGAGATCTGCGCAGAGAAATCTCTGCAAGCCCAAGTTAGTCTCTCTC 338
 Qy 61 CAAGGCAAGGAGCTGAGCATGAGAGCAAGAGCAAGCCAGCGTGGCCCTG 120
 Db 339 CAAGGCAAGGAGCTGAGCATGAGAGCAAGAGCAAGCCAGCGTGGCCCTG 398
 Qy 121 GGCAGTTTCCGCGAGGTGCCCCGCGAGCTGTGCTGAGACTGGGAGCAATTGAC 180
 Db 399 GGCAGTTTCCGCGAGGTGCCCCGCGAGCTGTGCTGAGACTGGGAGCAATTGAC 458
 Qy 181 ATGCTCTGAGATGAGAGCTGAGAGCTGCTGAGAGCTGAGAGCTGAGAGCTGAT 240
 Db 459 ATGCTCTGAGATGAGAGCTGAGAGCTGCTGAGAGCTGAGAGCTGAGAGCTGAT 518
 Qy 241 AACATCCCAAGGCTGACGTCGCAAGTCTCCATGAGGTGCTGATGAGGCTGAGC 300
 Db 519 AACATCCCAAGGCTGACGTCGCAAGTCTCCATGAGGTGCTGATGAGGCTGAGC 578
 Qy 301 AGGAGAAAGAGAGAGAACTGCTGTTTACTGAGGAACTCTGAGAGGAGG--CTTCTCTAT 359
 Db 579 AGGAGAAAGAGAGAGAACTGCTGTTTACTGAGGAACTCTGAGAGGAGG--CTTCTCTAT 638
 Qy 360 CCGGAGAG-AGCCAGACAGAGAGGCTTTACTCTGCTGAGT-CCGCTTACGCG-CCC 416
 Db 639 CCGGAGAGAGCAAGACAGAGAGGCTTTACTCTGCTGAGT-CCGCTTACGCG-CCC 698
 Qy 417 TGCATCTGAGG--ACCGATCAGACACTACAGG 447
 Db 699 TGCATCTGAGGACCGGATCAGACACTACAGG 731

RESULT 8

LOCUS B0053486 878 bp mRNA linear EST 29-MAR-2002
 DEFINITION AGENCOURT_6822017 NIH_MGC_106 Homo sapiens CDNA clone IMAGE:5935253
 5', mRNA sequence.
 ACCESSION B0053486

Db 721 TCCCGAGAGACAGACAGAGAGCTCTTA 752

RESULT 10
AL844311
LOCUS AL844311 597 bp mRNA linear EST 30-JUL-2002
DEFINITION AL844311 pool_YT_11b_v_SPD Homo sapiens cDNA, mRNA sequence.
ACCESSION AL844311
VERSION AL844311.1 GI:22019093
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 597)
Ashcroft,K., Bethel,G., Bye,J.M., Howell,G.R., Huckle,E.J. and
Sheridan,E.
Homo sapiens EST sequence
Unpublished (2002)
Contact: The Sanger Centre
The Sanger Centre
Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK
Email: humquery@sanger.ac.uk
Sanger Centre name: scd10816.154136A
Homo sapiens EST sequence. This sequence was generated as part of
The Wellcome Trust Sanger Institute program to identify and
annotate genes in the human genome. Incomplete or unconfirmed genes
are experimentally analysed using a variety of cDNA library
resources. This sequence was obtained from a PCR product generated
from a pool of up to 100,000 cDNA clones derived from
pool YT_11b v SPD cDNA library. Further information can be found at
<http://www.sanger.ac.uk/Teams/Teams6/>.
Location/Qualifiers
1..597
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="20"
/clone_11b="pool_YT_11b_v_SPD"
/note="Organ: breast; Vector: pZERO-1; Site: 1: SpH1;
Site 2: SpH1; Ductal carcinoma in situ, high-grade, comedo
from 41 yo female. Library constructed in the laboratory
of K. Polyak, M.D., Ph.D. (Dana-Farber Cancer Institute)."

FEATURES

source

BASE COUNT 130 a 186 c 157 g 124 t
ORIGIN

Query Match 40.8%; Score 319.4; DB 9; Length 597;
Best Local Similarity 99.7%; Pred. No. 2.8e-71;
Matches 320; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 463 AATGGCTGGCTGTAATCTCAACGCGCCCTCACTCCCTCACTCCAGGCCCTGTGTGAC 522
Db 13 AATGGCTGGCTGTAATCTCAACGCGCTCACTCCCTCACTCCAGGCCCTGTGTGAC 72
Qy 523 CATTACTGAGCTGGCGGATGATCTGCTGCTCACTCAAGAGCCCTGTGTGAC 582
Db 73 CATTACTGAGCTGGCGGATGATCTGCTGCTCACTCAAGAGCCCTGTGTGAC 122
Qy 553 AGGCGTGGCCCTCCCTGGAGAGATATACCCCTCACTGAGCTGGAGAGACCA 642
Db 133 AGGCGTGGCCCTCCCTGGAGAGATATACCCCTCACTGAGCTGGAGAGACCA 192
Qy 643 CTCAACTGGAAGAGCTGACAGCTCCCTCTGTTTCTGAAGTGCACAGAGGAGAG 702
Db 193 CTCAACTGGAAGAGCTGACAGCTCCCTCTGTTTCTGAAGTGCACAGAGGAGAG 252
Qy 703 TCTCTTCTGAGAGGCTCTGGGAGTCCCTCACTTCAATCAAGCTGAATGACAG 762
Db 253 TCTCTTCTGAGAGGCTCTGGGAGTCCCTCACTTCAATCAAGCTGAATGACAG 312
Qy 763 GCTGTCTCTTTGATGATGCC 783
Db 313 GCTGTCTCTTTGATGATGCC 333

RESULT 11

AL844307 614 bp mRNA linear EST 30-JUL-2002
LOCUS AL844307
DEFINITION AL844307 pool_AK_11b_v_SPD Homo sapiens cDNA, mRNA sequence.
ACCESSION AL844307
VERSION AL844307.1 GI:22019089
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 614)
Ashcroft,K., Bethel,G., Bye,J.M., Howell,G.R., Huckle,E.J. and
Sheridan,E.
Homo sapiens EST sequence
Unpublished (2002)
Contact: The Sanger Centre
The Sanger Centre
Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK
Email: humquery@sanger.ac.uk
Sanger Centre name: scd10816.154136A
Homo sapiens EST sequence. This sequence was generated as part of
The Wellcome Trust Sanger Institute program to identify and
annotate genes in the human genome. Incomplete or unconfirmed genes
are experimentally analysed using a variety of cDNA library
resources. This sequence was obtained from a PCR product generated
from a pool of up to 100,000 cDNA clones derived from
pool AK_11b v SPD cDNA library. Further information can be found at
<http://www.sanger.ac.uk/Teams/Teams6/>.
Location/Qualifiers
1..614
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="20"
/clone_11b="pool_AK_11b_v_SPD"
/note="Organ: breast; Vector: pZERO-1; Site: 1: SpH1;
Site 2: SpH1; Ductal carcinoma in situ, high-grade, comedo
from 41 yo female. Library constructed in the laboratory
of K. Polyak, M.D., Ph.D. (Dana-Farber Cancer Institute)."

FEATURES

source

BASE COUNT 134 a 188 c 164 g 128 t
ORIGIN

Query Match 40.2%; Score 314.4; DB 9; Length 614;
Best Local Similarity 99.7%; Pred. No. 5.5e-70;
Matches 315; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 468 CTGGCTGATCATCAACGCGCTCACTCCCTCACTCCAGGCCCTGTGTGACATTA 527
Db 1 CTGGCTGATCATCAACGCGCTCACTCCCTCACTCCAGGCCCTGTGTGACATTA 60
Qy 528 CTGTAGCTGGCGGATGATCTGCTGCTCACTCAAGAGCCCTGTGTGAC 587
Db 61 CTGTAGCTGGCGGATGATCTGCTGCTCACTCAAGAGCCCTGTGTGAC 120
Qy 588 TGGCCGCTCCCTGGAGAGATATACCCCTCACTGAGCTGGAGAGACCACTCA 647
Db 121 TGGCCGCTCCCTGGAGAGATATACCCCTCACTGAGCTGGAGAGACCACTCA 180
Qy 648 CTGGAAGAGCTGACAGCTCCCTCTGTTTCTGAAGTGCACAGAGGAGAGTCT 707
Db 181 CTGGAAGAGCTGACAGCTCCCTCTGTTTCTGAAGTGCACAGAGGAGAGTCT 240
Qy 708 TCTCACTGAGAGTCTCCGAGAGTCCCTCACTTCAATCAAGCTGAATGACAGGCTGT 767
Db 241 TCTCACTGAGAGTCTCCGAGAGTCCCTCACTTCAATCAAGCTGAATGACAGGCTGT 300
Qy 768 CTCTTTGATGATGCC 783
Db 301 CTCTTTGATGATGCC 316

RESULT 12

cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGACAGATCTCCAGTAAATTAATATCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pbluescript KS(+) after bulk excision from Lambda PLC I."

BASE COUNT 141 a 191 c 187 g 140 t 1 others
ORIGIN

Query Match 37.8%; Score 296.2; DB 10; Length 660;
Best Local Similarity 82.9%; Pred. No. 2.6e-65;
Matches 350; Conservative 0; Mismatches 69; Indels 3; Gaps 1;

1 ATGGGAAGTCTCCAGACAGAAATATCTGCCAACCCAGCTTGAATCTCTCTC 60
241 ATGGGAAGTCTCCAGACAGAAATATCTGCCAACCCAGCTTCTCTCTCTG 297
61 CAAGCCAGGACCTGTGACCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
298 CAGACACAGAGAACCGGTGTCCATGCAACAGAGAGAGAGAGAGAGAGAGAG 357
121 GCGATTTTCCCGGAGGTGCGCCGCGGAGCTGCTGAGATTCGGGAGAGCAAT 180
358 GCGATTTTCCCGGAGGTGCGCCGCGGAGCTGCTGAGATTCGGGAGAGCAAT 417
181 ATGCTCTGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 240
418 ATCATCTCTGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 477
241 AACATCCCGAGAGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
478 CACATGCCAGTGTATGTGCTTAAGTGTGCTTAAGTGTGCTTAAGTGTGCT 537
301 AGGAG 360
538 CCGAG 597
361 CCGAG 420
598 CCGAG 657
421 TC 422
658 TC 659

RESULT 14
AA959151 377 bp mRNA linear EST 08-MAY-1998
LOCUS v51906.x1 Soares, thymus 2nbmt Mus musculus cDNA clone
IMAGE:1330042 5' similar to TR:Q13239 Q13239 PUTATIVE SRC-LIKE
ADAPTER PROTEIN ; mRNA sequence.
AA959151
AA959151.1 GI:3124344
EST.
VERSION
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE
AUTHORS Mammalia; Eutheria; Rodentia; Scurionath; Muridae; Murinae; Mus.
1 (bases 1 to 377)
Marr, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Lee, M., Martin, J., Morris, M.,
Thelning, B., Wylie, T., Lemon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800

Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:689586
Seq primer: -28m3 rev2 ET from Amersham.
Location/Qualifiers

FEATURES

SOURCE

1..377
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1330042"
/clone_1db="Soares_thymus_2nbmt"
/sex="male"
/tissue_type="Thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Vector: pT773D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAGAGTGGAGCGCGCGGCTTTTCTTTTCTTTTCTTTTCTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT773 vector. RNA
provided by Dr. Betrand Jordan. Library went through two
rounds of normalization, and was constructed by Bento
Soares and M. Fatima Bonaldi."

BASE COUNT

ORIGIN

Query Match 36.2%; Score 283.6; DB 9; Length 377;
Best Local Similarity 86.6%; Pred. No. 3.5e-62;
Matches 324; Conservative 0; Mismatches 49; Indels 1; Gaps 1;

115 GCCCTGGGCAATTTCCCGAGGTGCGCCGCGAGAGCTGCTGAGACTCGGAGAGCA 174
5 GGCCTGGGCAATTTCCCGAGGTGCGCCGCGAGAGCTGCTGAGACTCGGAGAGCG 64
175 TTGACCATCTGCTCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 234
65 CTGACCATCTGCTCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 124
235 GAGTATTAACATCCCAAGCTTCACTGGCCAAAGTCTCCATGAGTGTGATGAGGC 294
125 GAGTATTAACATCCCAAGCTTCACTGGCCAAAGTCTCCATGAGTGTGATGAGGC 184
295 CTGAG 354
185 CTGAG 243
355 CTGATCCGGAG 414
244 CTGATCCGGAG 303
415 CTGATCCGGAG 474
304 CTGATCCGGAG 363
475 TACATCTACCGCG 488
364 TACATCTACCGCG 377

RESULT 15
BB619854 627 bp mRNA linear EST 31-AUG-2001
LOCUS BB619854 RIKEN full-length enriched, adult male thymus Mus musculus
DEFINITION BB619854 RIKEN full-length enriched, adult male thymus Mus musculus
ACCESSION BB619854
KEYWORDS BB619854.1 GI:15396929
VERSION
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE
AUTHORSTITLE
JOURNAL
COMMENTFEATURES
source

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 627)

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hangaki, T., Hara, A., Hiramoto, K., Horii, F., Iehli, Y., Ito, M., Kawai, J., Kono, H., Konda, M., Koye, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ono, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M., and Hayashizaki, Y.

RIKEN Mouse ESTs (Arakawa, T., et al. 2001)

Unpublished (2001)

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Sphiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

wegli, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y., and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Yamanaka, I., Kiyosawa, H., Kondo, S., Saito, T., Shinagawa, A., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shiraki, K., Arakawa, T., Iehli, Y., and Hayashizaki, Y.

Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct. Funct. Genomics 2 pre, U72-U86 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

e mouse tissues.

Location/Qualifiers

1. 627
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="5830437K10"
/clone_id="RIKEN full-length enriched, adult male thymus"
/sex="male"
/tissue_type="thymus"
/dev_stage="adult"
/lab_host="MDH108"
/note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGATTCGAGATTCGATTAATTAATTAATCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision

BASE COUNT 138 a 179 c 180 g 130 t
ORIGIN
Query Match 35.2%; Score 275.6; DB 10; Length 627;
Best Local Similarity 82.1%; Pred. No. 4.9e-60;
Matches 330; Conservative 0; Mismatches 69; Indels 3; Gaps 1;
From Lambda FLC I. Cloning sites, 5' end: SalI, 3' end: BamHI.

1 ATGGAGTGTGCGCCAGAGAAATCTTGGCCAGCCAGCTTGAAGTCTTGTGC 60
229 ATGGAGTGTGCGCCAGAGAAATCTTGGCCAGCCAGCTTGAAGTCTTGTGC 285
61 CAAGGCGAGGAGCTGTGACCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
286 CCAAGCAG 345
121 GGCAGTTTCCGCGAG 180
346 GGCAGTTTCCGCGAG 405
181 ATCGTCTGTGAT 240
406 ATCATCTGTGAT 465
241 AACATCCCGAGCGTCCAGTGGCCAAAGTCTCCATGAGTGGTGTGATGAGAGAG 300
466 CACATGCCAGTGTGTATGTGTGCTTAAAGTCCGCCAGGAGTGTGTATGAGAGAG 525
301 AGGAG 360
526 CCGAG 585
361 CCGAG 402
586 CCGAG 627

Search completed: March 30, 2003, 02:49:17
Job time : 1547.85 secs

